

George, Ann Cathrice; Robitzsch, Alexander; Schreiner, Claudia  
**Eine Diskussionsgrundlage zur Weiterentwicklung von Rückmeldungen aus  
standardisierten Kompetenzmessungen am Beispiel Mathematik**

George, Ann Cathrice [Hrsg.]; Schreiner, Claudia [Hrsg.]; Wiesner, Christian [Hrsg.]; Pointinger, Martin [Hrsg.]; Pacher, Katrin [Hrsg.]: *Kompetenzmessungen im österreichischen Schulsystem: Analysen, Methoden & Perspektiven*. [1. Auflage]. Münster ; New York : Waxmann 2019, S. 225-238. - (*Kompetenzmessungen im österreichischen Schulsystem: Analysen, Methoden & Perspektiven*; 1)



Quellenangabe/ Reference:

George, Ann Cathrice; Robitzsch, Alexander; Schreiner, Claudia: Eine Diskussionsgrundlage zur Weiterentwicklung von Rückmeldungen aus standardisierten Kompetenzmessungen am Beispiel Mathematik - In: George, Ann Cathrice [Hrsg.]; Schreiner, Claudia [Hrsg.]; Wiesner, Christian [Hrsg.]; Pointinger, Martin [Hrsg.]; Pacher, Katrin [Hrsg.]: *Kompetenzmessungen im österreichischen Schulsystem: Analysen, Methoden & Perspektiven*. [1. Auflage]. Münster ; New York : Waxmann 2019, S. 225-238 - URN: urn:nbn:de:0111-pedocs-178108 - DOI: 10.25656/01:17810

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-178108>

<https://doi.org/10.25656/01:17810>

in Kooperation mit / in cooperation with:



**WAXMANN**  
[www.waxmann.com](http://www.waxmann.com)

<http://www.waxmann.com>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

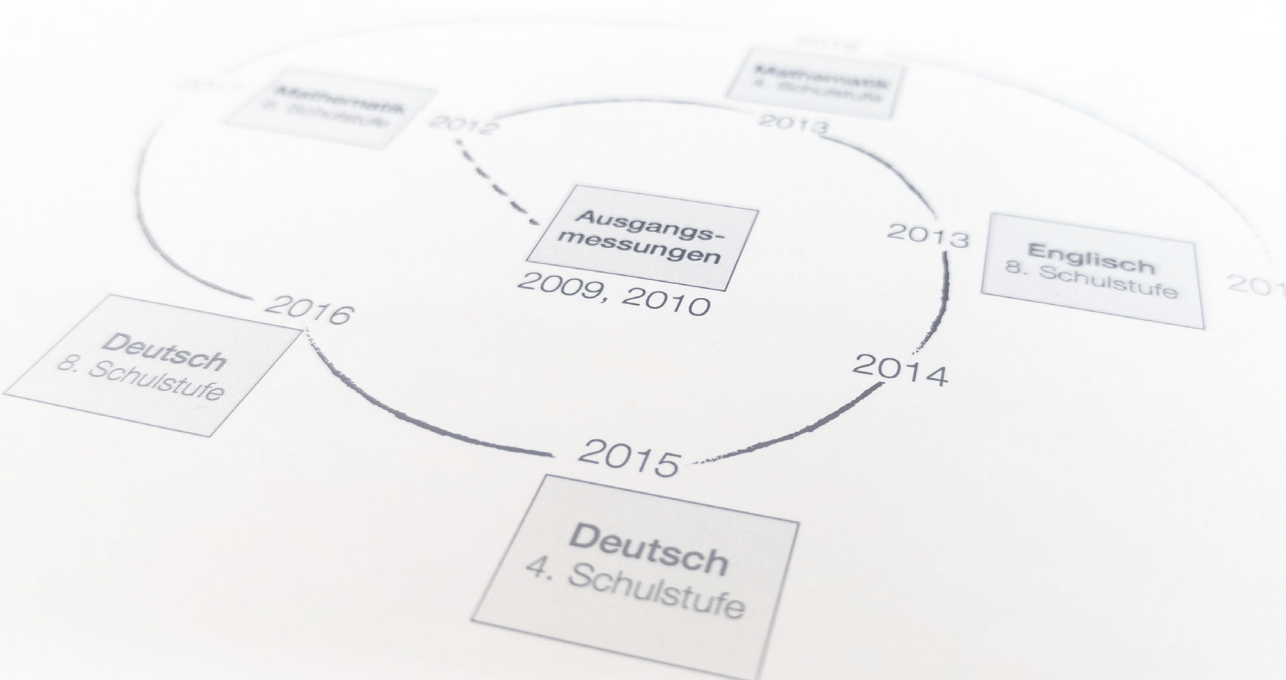
peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)



Ann Cathrice George, Claudia Schreiner, Christian Wiesner,  
Martin Pointinger, Katrin Pacher (Hrsg.)

# Fünf Jahre flächendeckende Bildungsstandardüberprüfungen in Österreich

Vertiefende Analysen zum Zyklus 2012 bis 2016



WAXMANN



Bundesinstitut  
**bifie**

# Kompetenzmessungen im österreichischen Schulsystem: Analysen, Methoden & Perspektiven

herausgegeben vom  
Bundesinstitut für Bildungsforschung,  
Innovation und Entwicklung des  
österreichischen Schulwesens (BIFIE)

Band 1

Ann Cathrice George, Claudia Schreiner,  
Christian Wiesner, Martin Pointinger, Katrin Pacher (Hrsg.)

# Fünf Jahre flächendeckende Bildungsstandardüberprüfungen in Österreich

Vertiefende Analysen zum Zyklus 2012 bis 2016



Waxmann 2019  
Münster • New York

### **Bibliografische Informationen der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

### **Kompetenzmessungen im österreichischen Schulsystem: Analysen, Methoden & Perspektiven, Bd. 1**

ISSN 2628-9598

Print-ISBN 978-3-8309-3925-2

© Waxmann Verlag GmbH, 2019

[www.waxmann.com](http://www.waxmann.com)

[info@waxmann.com](mailto:info@waxmann.com)

Umschlaggestaltung: Hannes Kaschnig-Löbel, Salzburg

Satz: Stoddart Satz- und Layoutservice, Münster

Druck: CPI books GmbH, Leck

Gedruckt auf alterungsbeständigem Papier,  
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.

Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

# Inhalt

Vorwort der Direktorin des BIFIE .....	7
--	---

## Einleitung

<i>Claudia Schreiner und Christian Wiesner</i>	
<b>1</b> Die Überprüfung der Bildungsstandards in Österreich: der erste Zyklus als Meilenstein für die Schul- und Unterrichtsentwicklung – eine gelungene Innovation im österreichischen Schulsystem .....	13
<i>Maria Neubacher, Michael Ober, Christian Wimmer und Martina Hartl</i>	
<b>2</b> Die Kompetenzen der Schüler/innen in der Zusammenschau aller Standardüberprüfungen des ersten Zyklus .....	55

## Fokussierte Analysen auf Ebene der Schulen und Klassen

<i>Roman Freunberger, Lisa Mayrhofer und Judith Sauerwein</i>	
<b>3</b> Die Situation von Klein- und Kleinstschulen in der Volksschule .....	89
<i>Ann Cathrice George und Susanne Schwab</i>	
<b>4</b> Österreichs Integrationsklassen: Kompetenzdefizite durch soziale Benachteiligung? Ein Vergleich zwischen Integrations- und Regelklassen .....	103
<i>Claudia Schreiner, Christian Wiesner, Thomas Kiefer, Christoph Helm, Mishela Ivanova, David Kemethofer, Marcel Illetschko, Margit Freller-Töglhofer und Daniel Paasch</i>	
<b>5</b> Merkmale des fachlichen Unterrichts und Schülerkompetenzen .....	115

## Fokussierte Analysen auf Ebene der Schüler/innen

<i>Claudia Schreiner, Burkhard Gniewosz, Christian Wiesner, Alexander Steiger, Andrea Kulmhofer-Bommer und Maximilian Egger</i>	
<b>6</b> Einstellung der Schüler/innen zum Fach und zum Lernen: Freude am Fach, fachbezogenes Selbstkonzept und ihr Zusammenhang mit den fachlichen Leistungen .....	139

	<i>Daniel Paasch, Christine Schmid, Andrea Kallinger-Aufner und Robert Knollmüller</i>	
<b>7</b>	Noten und Kompetenzen in verschiedenen Fächern, Schulstufen und Schulformen.....	161
	<i>Simone Breit, Barbara Herzog-Punzenberger, Silvia Salchegger und Philipp Schnell</i>	
<b>8</b>	Mehrsprachige Schüler/innen am Ende der 8. Schulstufe: Kompetenzen und familiäres Sprachumfeld .....	179
	<i>Konrad Oberwimmer, Norbert Lachmayr und Silke Luttenberger</i>	
<b>9</b>	Bildungsaspiration und Berufsberatungsmaßnahmen.....	199

## **Ausblick und Diskussion von Ergebnissen aus dem ersten Zyklus**

	<i>Ann Cathrice George, Alexander Robitzsch und Claudia Schreiner</i>	
<b>10</b>	Eine Diskussionsgrundlage zur Weiterentwicklung von Rückmeldungen aus standardisierten Kompetenzmessungen am Beispiel Mathematik.....	225
	<i>Christian Wiesner, Claudia Schreiner, Daniel Paasch, Simone Breit und Roman Freunberger</i>	
<b>11</b>	Der kompetenzorientierte Unterricht in Mathematik in der österreichischen Volksschule aus Sicht der Schüler/innen: Impulse für eine reflexive Unterrichtsentwicklung und -forschung.....	239
	Autorinnen und Autoren.....	259

Die Analysen im vorliegenden Bericht wurden an Daten durchgeführt, die weitestgehend in der Forschungsdatenbibliothek (FDB) des BIFIE dokumentiert, archiviert und Wissenschaftlerinnen und Wissenschaftlern über ein Antragsverfahren zugänglich gemacht werden.

Neben Beschreibungen zum Antragsverfahren werden auf der Webseite der FDB (<https://www.bifie.at/fdb>) die Daten zur Vorbereitung von Forschungsanträgen transparent gemacht. Für jeden Forschungsdatensatz finden sich dort Kontextfragebögen, Codebücher und Almanache (univariate deskriptive Statistiken der verfügbaren Variablen).

Die Wahrung der Anonymität der Untersuchungseinheiten und der Grundsatz der Datenminimierung sind im BIFIE-Gesetz<sup>1</sup> bzw. in der DSGVO<sup>2</sup> einerseits gesetzlich verankert und tragen andererseits einen wesentlichen Anteil zur Akzeptanz der Erhebungen bei. Die Weiterverarbeitung der BIST-Daten im Rahmen wissenschaftlicher Fragestellungen ist unter Einhaltung dieser Bedingungen nicht nur gestattet,<sup>3</sup> sondern sogar gewünscht.

Die FDB gewährleistet die Anonymität der Untersuchungseinheiten durch ihr Sicherheitskonzept, indem klare Verantwortlichkeiten festgelegt werden, sowie durch die Bereitstellung von technisch anonymisierten Forschungsdaten.

Mit dem R-Paket BIFIEsurvey (BIFIE, 2018; siehe auch Bruneforth, Oberwimmer & Robitzsch, 2016) und den BIFIE-SPSS-Makros stellt die FDB Werkzeuge zur Verfügung, die speziell für die Analysen von FDB-Forschungsdaten entwickelt wurden.

<sup>1</sup>BGBl. I Nr. 25/2008 i.d.F. I 32/2018, § 3 Abs. 1 Z. 5, § 6 Abs. 3 bzw. § 7b Abs. 1.

<sup>2</sup>Art. 5 lit c VO (EU) 2016/679

<sup>3</sup>Art. 5 lit b und 89 VO (EU) 2016/679 sowie § 2d Abs 4 FOG

#### Literatur

BIFIE (2018). BIFIEsurvey: Tools for survey statistics in educational assessment. R package version 3.0-14 [Software]. Verfügbar unter <https://CRAN.R-project.org/package=BIFIEsurvey>

BIFIE (2017). SPSS-Makros. Version v1.6 [Software]. Verfügbar unter <https://www.bifie.at/fdb>

Bruneforth, M., Oberwimmer, K. & Robitzsch, A. (2016). Reporting und Analysen. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R. Methodische Grundlagen der österreichischen Bildungsstandardüberprüfungen*. (S. 333–362). Wien: facultas.

Abbildung 1: Informationen zur Forschungsdatenbibliothek des BIFIE.

#### Information für in Grafiken dargestellte Ergebnisse:

In der Regel wurden die angegebenen Kennwerte (Punktwerte, Prozentangaben etc.) in diesem Bericht unter Berücksichtigung entsprechender Nachkommastellen berechnet und dann auf ganze Zahlen gerundet. Daher kann es vorkommen, dass die Summe der gerundeten Prozentangaben nicht exakt 100 ergibt oder Summen von Werten inkonsistent erscheinen mögen.



## **10 Eine Diskussionsgrundlage zur Weiterentwicklung von Rückmeldungen aus standardisierten Kompetenzmessungen am Beispiel Mathematik**

### **Zusammenfassung**

Weil die Ergebnissrückmeldungen für das System und an Lehrer/innen über ihre Unterrichtsgruppen unterschiedliche Ziele verfolgen, unterscheiden sich auch die Anforderungen an die Rückmeldungen auf diesen beiden Ebenen. Basierend auf dieser Annahme werden im vorliegenden Kapitel Entwicklungsdesiderate in der Rückmeldung von Mathematikkompetenzen auf Unterrichtsebene herausgestellt und ein Vorschlag für eine Weiterentwicklung der Rückmeldungen eingebracht. Das Kapitel schließt mit einer Diskussion der Vor- und Nachteile der aktuellen Rückmeldeform im Vergleich zu der neu vorgeschlagenen.

### **10.1 Aktuelle Ergebnissrückmeldung aus der Bildungsstandardüberprüfung im Fach Mathematik**

Im ersten Zyklus der Bildungsstandardüberprüfungen (BIST-Ü) wurden die Kompetenzen der Schüler/innen in Mathematik in der achten Schulstufe im Jahr 2012 (Schreiner & Breit, 2012) und in der vierten Schulstufe im Jahr 2013 (Schreiner & Breit, 2014) erhoben. Beide Überprüfungen wurden als Vollerhebungen durchgeführt, d. h. alle Schüler/innen, die in den jeweiligen Jahren die betreffende Schulstufe in Österreich besuchten, wurden getestet.<sup>1</sup> Eine Grundlage für den Aufbau der eingesetzten Tests und der daraus resultierenden Ergebnissrückmeldungen bilden Kompetenzmodelle für die jeweilige Schulstufe. Die Erläuterungen in diesem Kapitel werden im Weiteren anhand der Ergebnisse für die Sekundarstufe diskutiert, lassen sich aber auf die Volksschule übertragen.

---

<sup>1</sup> Nach der Verordnung der Bildungsstandards BGBl. II Nr. 1/2009 sind einige wenige Schüler/innen von den Überprüfungen ausgenommen. Darunter fallen außerordentliche Schüler/innen, Schüler/innen mit sonderpädagogischem Förderbedarf (wenn sie im Testfach nicht nach dem Lehrplan der Volksschule oder der Neuen Mittelschule beurteilt werden) und Schüler/innen mit Körper- und Sinnesbehinderungen.

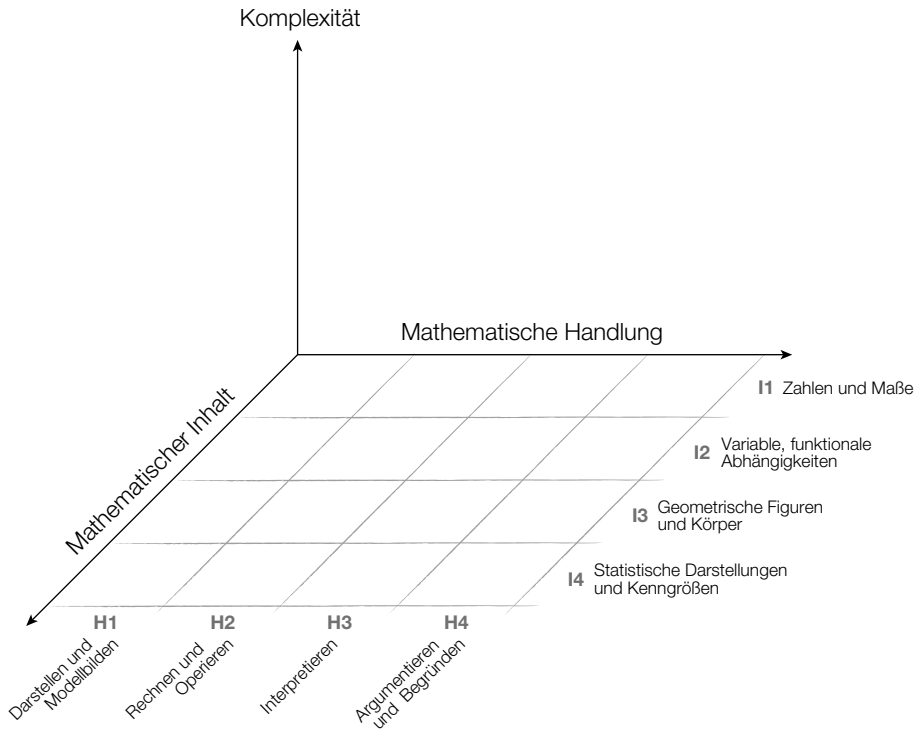


Abbildung 1: Kompetenzmodell Mathematik für die Sekundarstufe 1 (in Anlehnung an Schreiner et al., 2018, S. 16).

Das Kompetenzmodell für Mathematik auf der Sekundarstufe 1 (vgl. Abbildung 1) sieht eine Aufgliederung der mathematischen Kompetenz in Inhaltsbereiche, Handlungsbereiche und Komplexitätsgrade vor. Die vier Inhaltsbereiche (auch inhaltliche Kompetenzbereiche) umfassen *Zahlen und Maße*, *Variable und funktionale Abhängigkeiten*, *Geometrische Figuren und Körper* sowie *Statistische Darstellungen und Kenngrößen*. Die Handlungsbereiche umfassen *Darstellen und Modellbilden*, *Rechnen und Operieren*, *Interpretieren* sowie *Argumentieren und Begründen*. Die Komplexitätsgrade beschreiben das Niveau, auf dem Fakten, Zusammenhänge, Darstellungen und Handlungen miteinander in Verbindung gebracht werden müssen. Dieser Struktur folgend werden bei der Ergebnismrückmeldung neben einem Gesamtwert der mathematischen Kompetenz auch Werte der durch die Schüler/innen beherrschten Kompetenzen in den vier Inhaltsbereichen und den vier Handlungsbereichen rückgemeldet.<sup>2</sup> Diese detaillierte Rückmeldung<sup>3</sup> ist unter anderem auch dadurch möglich, dass bereits bei der Konstruktion der Testaufgaben darauf geachtet wird, dass zur Lösung jedes Items die Kompetenzen aus nur einem der vier inhaltlichen Bereiche und nur einem der vier Handlungsbereiche ausschlaggebend sind (vgl. auch Itzlinger-Bruneforth, Kuhn & Kiefer, 2016).

<sup>2</sup> Die Komplexitätsgrade finden in der Ergebnismrückmeldung keine Beachtung, da sich diese empirisch nicht abgrenzen lassen.

<sup>3</sup> Die Ergebnismrückmeldung beinhaltet ein deutlich breiteres Spektrum als nur die Präsentation der Kompetenzwerte in den Handlungs- und Inhaltsbereichen. Auf der Darstellung der Kompetenzwerte liegt allerdings der Hauptfokus dieses Kapitels.

Die Ergebnisse der Bildungsstandardüberprüfungen werden in einem detaillierten Berichterstattungs- und Rückmeldesystem dokumentiert, welches alle Ebenen umfasst (vgl. z. B. Schreiner & Breit, 2014, S. 13 f.). So werden auf oberster Ebene die Ergebnisse für die System- bzw. Bundesebene sowie für die Bundesländer aufgearbeitet. Diese Ergebnisse werden in Form von Berichten in öffentlich zugänglichen Publikationen bereitgestellt. Weiters erhalten jede überprüfte Schule, die Lehrpersonen von jeder überprüften Unterrichtsgruppe sowie jede überprüfte Schülerin und jeder überprüfte Schüler eine individuelle (nicht öffentliche) Rückmeldung. Zudem werden die Informationen aus den Schulberichten zusammengefasst und der Schulaufsicht als empirische Grundlage zur Begleitung von schulischen Qualitätsentwicklungsprozessen zur Verfügung gestellt. Die Form der Rückmeldungen bzw. der Rückmeldedaten auf den verschiedenen Ebenen ist an die jeweiligen Adressaten angepasst.

Zur Erstellung der Rückmeldungen auf allen Ebenen werden die Antworten der Schüler/innen zu den Testaufgaben codiert und anschließend mit ein- und mehrdimensionalen Item-Response-Modellen (IRT-Modellen) skaliert (Trendtel, Pham & Yanagida, 2016). Die Ergebnispräsentation erfolgt auf einer 500er-Metrik. Abbildungen 2 bis 4 zeigen durch dieses Vorgehen erzeugte Rückmeldungen<sup>4</sup> in den vier Inhaltsbereichen und den vier Handlungsbereichen. Die sich vom Aufbau ähnelnden Rückmeldungen sind jeweils auf den Adressaten bezogen, d. h. Abbildung 2 enthält Informationen über ganz Österreich (Bundesebene), Abbildung 3 fasst die Ergebnisse einer Unterrichtsgruppe zusammen (Gruppenebene) und Abbildung 4 enthält Ergebnisse für eine überprüfte Schülerin bzw. einen überprüften Schüler (Individualebene).

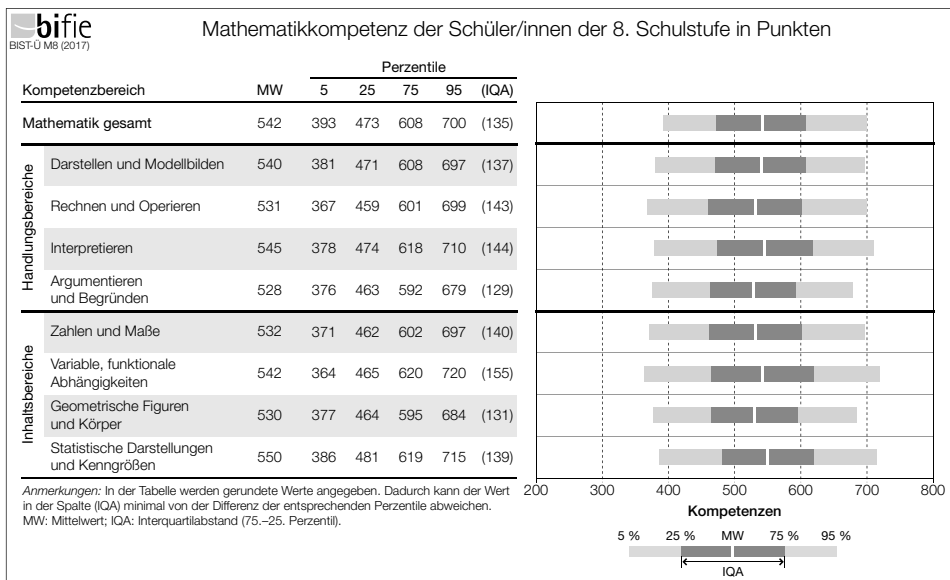


Abbildung 2: Rückmeldung der vier Inhaltsbereiche (unten) und Handlungsbereiche (oben) im Bundesergebnisbericht der BIST-Ü M8 2017.

4 Wir zeigen hier die neueste Form der Rückmeldung von Mathematikkompetenzen nach der Bildungsstandardüberprüfung M8 2017 (Schreiner et al., 2018).

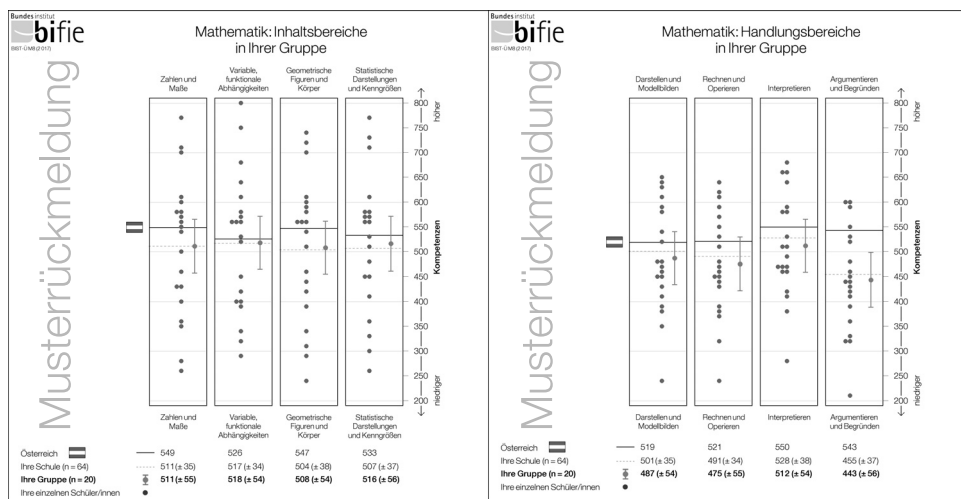


Abbildung 3: Rückmeldung der vier Inhaltsbereiche (links) und vier Handlungsbereiche (rechts) für eine Unterrichtsgruppe bei der BIST-Ü M8 2017.

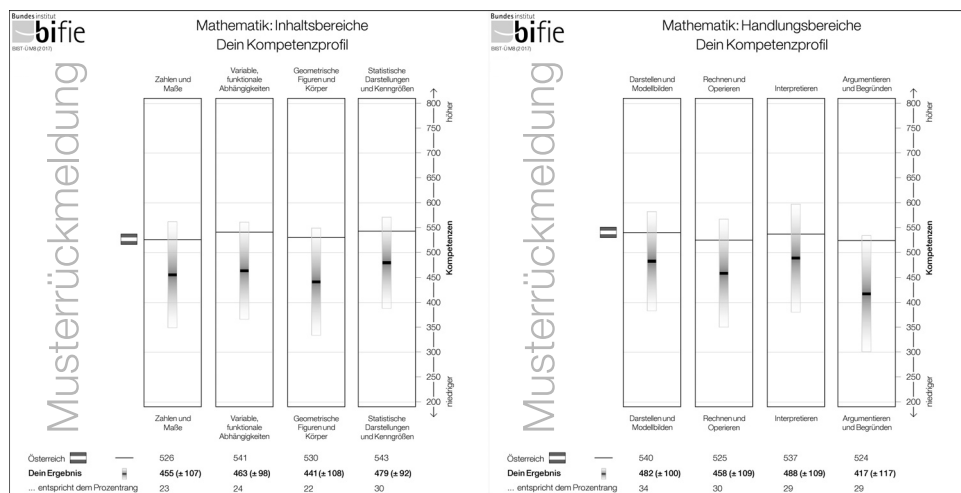


Abbildung 4: Rückmeldung der vier Inhaltsbereiche (links) und vier Handlungsbereiche (rechts) für eine Schülerin bzw. einen Schüler in der BIST-Ü M8 2017.

Auf der Systemebene (Abbildung 2) ist für jeden der acht Kompetenzbereiche der durchschnittliche Kompetenzwert österreichischer Schüler/innen zu entnehmen. Beim Vergleich dieser Kompetenzwerte mit dem in einer Ersterhebung 2009 (Baseline-Studie) definierten Kompetenzwert von 500 kann für jeden Kompetenzbereich der Kompetenzzuwachs zwischen 2009 und 2017 abgeleitet werden. So erfolgten in Mathematik in der Sekundarstufe zwischen 2009 und 2017 die größten Zuwächse im inhaltlichen Bereich „Statistische Darstellungen und Kenngrößen“ (+50 Punkte) und im Handlungsbereich „Interpretieren“ (+45 Punkte). Aufgrund des Erhebungsdesigns

bedeuten unterschiedliche durchschnittliche Kompetenzwerte in den Inhalts- und Handlungsbereichen unterschiedlich starke Entwicklungen auf Bundesebene seit der Ausgangsmessung. Da die Kompetenzbereiche nicht auf einer gemeinsamen Metrik verankert sind, lassen sich direkte Vergleiche zwischen den Kompetenzbereichen nur in dem Sinne des Zuwachses seit der Ausgangsmessung interpretieren.

Auf Ebene der Unterrichtsruppen (Abbildung 3) kann die jeweils zuständige Lehrkraft den mittleren Kompetenzwert ihrer Gruppe mit den mittleren Kompetenzwerten aller österreichischen Schüler/innen sowie aller Schüler/innen der eigenen Schule vergleichen. Daraus kann die Lehrkraft entnehmen, in welchen Kompetenzbereichen ihre Unterrichtsruppe im Vergleich zum österreichischen Mittel bzw. im Vergleich zur gesamten Schule besser bzw. schlechter abgeschnitten hat. Ein Vergleich der Unterrichtsruppenmittelwerte zwischen den Kompetenzbereichen ist auch auf dieser Ebene schwierig. Als Abhilfe bieten die Rückmeldungen ein (hier nicht gezeigtes) Referenzprofil zur Bestimmung relativer Stärken bzw. Schwächen der Unterrichtsruppe.

Sehr ähnliche Informationen zu jenen auf der Unterrichtsruppenebene finden sich auch in der Individualrückmeldung für die einzelne Schülerin bzw. den einzelnen Schüler. Abbildung 4 enthält für jeden Kompetenzbereich den von der Schülerin/dem Schüler erreichten Kompetenzwert. Die Schülerin/der Schüler kann in jedem Kompetenzbereich ihren/seinen erreichten Wert mit dem österreichischen Mittelwert vergleichen. Ein direkter Vergleich zwischen den Werten der Kompetenzbereiche ist auch auf der Individualebene nicht sinnvoll zu interpretieren.

Die Rückmeldungen umfassen auf allen Ebenen weitere Aspekte, die für die folgenden Betrachtungen nicht von Bedeutung sind und deshalb nicht weiter ausgeführt werden.<sup>5</sup> Erwähnt werden sollte jedoch die kriteriale Rückmeldung. Für die Kompetenz in Mathematik wurden auf Basis der in den Bildungsstandards beschriebenen anzustrebenden Kompetenzen Kompetenzstufen entwickelt, welche durch ein Standardsetting mit der kontinuierlichen Kompetenzskala verknüpft wurden (vgl. Luger-Bazinger, Freunberger & Itzlinger-Bruneforth, 2016). Dieses Vorgehen erlaubt, den Punktwert auf der 500er-Skala inhaltlich einzuordnen (Bildungsstandards nicht erreicht, teilweise erreicht, erreicht oder übertroffen). In der Rückmeldung auf Unterrichtsruppenebene ist die prozentuelle Verteilung der Schüler/innen auf die vier Kompetenzstufen sowie auf Individualebene die Zugehörigkeit einer Schülerin/eines Schülers zu einer Stufe enthalten. Allerdings ist die Zuordnung zu einer Stufe, insbesondere auf Individualebene, durch die Größe der Standardfehler mit entsprechender Unsicherheit behaftet. Für die einzelnen Handlungs- und Inhaltsbereiche liegen keine kriterialen Rückmeldungen vor.

---

<sup>5</sup> Musterrückmeldungen für alle Ebenen können unter <https://www.bifie.at/material/ueberpruefung-der-bildungsstandards/musterrueckmeldungen/> eingesehen werden.

## 10.2 Unterschiedliche Ziele von Ergebnismeldungen auf Ebene des Systems und der Unterrichtsgruppen

Nach Pellegrino, Chudowsky und Glaser (2001) sollten bei jeder Kompetenzmessung die drei Komponenten Kognition (zugrundeliegendes domänenspezifisches Kompetenzmodell), Observation (eingesetzte Testaufgaben zur Kompetenzmessung) und Interpretation (statistisches Modell zur Auswertung der Aufgaben und Methoden zur Interpretation der Ergebnisse) aufeinander abgestimmt werden (Abbildung 5). Ändert sich eine der drei Komponenten, so sollte überprüft werden, ob die verbleibenden zwei Komponenten zu der veränderten weiterhin stimmig sind.

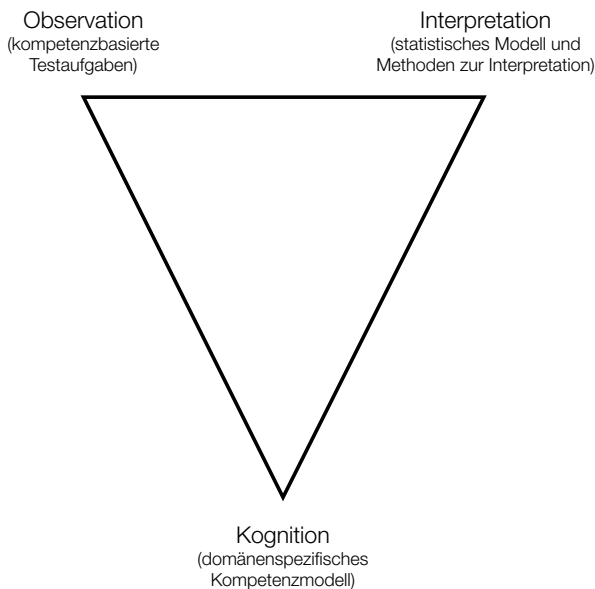


Abbildung 5: „Assessment Triangle“ mit den Komponenten Kognition, Observation und Interpretation (vgl. Pellegrino, Chudowsky und Glaser, 2001, S. 44).

Ziel der Systemrückmeldung (im Assessment Triangle: Ziel der Interpretation) ist, politischen Entscheidungsträgern grundlegende Informationen über das Bildungssystem zur Verfügung zu stellen. Basierend auf diesen Informationen können beispielsweise bundesweite Fördermaßnahmen für Schülergruppen diskutiert, die finanzielle Ressourcenplanung angepasst oder die Lehreraus- und -fortbildung weiterentwickelt werden. Durch die Ausrichtung der Kompetenzmessung auf die vier Inhalts- und die vier Handlungsbereiche (im Assessment Triangle: Kognition; vgl. Kompetenzmodell in Abbildung 1) können die Testaufgaben entsprechend entwickelt und die Tests entsprechend zusammengestellt werden (im Assessment Triangle: Observation). Zur Bestimmung der Kompetenzwerte in den Inhaltsbereichen werden die Testaufgaben über die Handlungsbereiche hinweg zusammengefasst. Es bilden also beispielsweise die Schülerantworten zu allen Aufgaben, welche dem inhaltlichen Be-

reich *Zahlen und Maße* zugeordnet sind, die Berechnungsgrundlage für den inhaltlichen Kompetenzwert des Inhaltsbereichs *Zahlen und Maße*. Dabei spielt es keine Rolle, welchem Handlungsbereich (*Darstellen und Modellbilden*, *Rechnen und Operieren*, *Interpretieren* sowie *Argumentieren und Begründen*) die Aufgaben zugeordnet sind. Analog erfolgen die Berechnungen für die Kompetenzwerte in den Handlungsbereichen durch Zusammenfassung aller Testaufgaben zu jeweils einem Handlungsbereich über alle inhaltlichen Bereiche hinweg. Durch die Skalierung der Schülerantworten in jedem der acht Kompetenzbereiche auf der 500er-Metrik und den Link zur Baseline 2009 (im Assessment Triangle: Interpretation, statistisches Modell) wird auf Systemebene eine Entwicklungsperspektive erreicht. Das Modell beschreibt in den acht Kompetenzbereichen Veränderungen im System zwischen den Messzeitpunkten, die dem Ziel der Interpretation auf Systemebene zweckdienlich sind.

Rückmeldungen an die Fachlehrer/innen auf Unterrichtsebene (im Assessment Triangle: Interpretation) haben ein anderes Ziel. Sie sollen den Lehrpersonen als Reflexionsgrundlage für eine fundierte und langfristige Unterrichtsentwicklung dienen (Wiesner et al., 2016). Basierend auf den empirischen Informationen aus der Rückmeldung können Lehrkräfte mithilfe ihres pädagogischen Fachwissens beispielsweise blinde Flecken in der Behandlung von Inhaltsbereichen erkennen oder Stärken in bestimmten Handlungsbereichen vor dem Hintergrund gewählter Unterrichtsmethoden reflektieren. Der Rückmeldung zugrunde liegen wiederum die vier Inhalts- und die vier Handlungsbereiche (im Assessment Triangle: Kognition). Durch einen dem Vorgehen auf der Systemebene analogen Zugang werden alle Testitems, die einem inhaltlichen oder einem Handlungsbereich zugeordnet sind (Assessment Triangle: Observation), skaliert und anschließend auf 500er-Metriken in jedem Kompetenzbereich präsentiert (im Assessment Triangle: Interpretation, statistisches Modell).

Die Rückmeldung in den acht Kompetenzbereichen kann allerdings als Diskrepanz zur tatsächlichen Unterrichtspraxis aufgefasst werden: Dem anfangs vorgestellten Kompetenzmodell und den zugrundeliegenden didaktischen Überlegungen zufolge findet Unterricht nicht getrennt nach Inhalts- bzw. Handlungsbereichen statt, sondern erfolgt durch die Kombination jeweils beider Kompetenzfacetten (sog. Kompetenzknoten; siehe auch Blum, 2012). Zum Ziel der Unterrichtsentwicklung wäre auf Unterrichtsebene eine Rückmeldung in den Kompetenzknoten wünschenswert, um detaillierte Informationen mit direkterem Bezug zum Unterricht zu gewinnen.<sup>6</sup>

Bezogen auf die Rückmeldung kann weiters hinterfragt werden (im Assessment Triangle: Interpretation), ob zum Ziel der Unterrichtsentwicklung ein normorientierter Vergleich der Unterrichtsgruppe nach außen (d.h. z.B. zum mittleren Wert der Kompetenzbeherrschung in Österreich, zum Mittelwert der Schule oder zu vergleichbaren Unterrichtsgruppen) eine relevante Information für die Fachlehrer/innen

6 Eine solche Form der Rückmeldung auf den Kompetenzknoten mittels IRT-Analysen könnte bei der Angabe von Mittelwerten und Standardabweichungen auf Gruppen- und Individualebene zu nicht gewünschter statistischer Unsicherheit führen. Um allerdings ein konsistentes Rückmeldeschema zwischen den Ebenen (System, Gruppe, Individual) beizubehalten, wurde in der aktuellen Rückmeldung auch auf der Bundesebene auf eine solche Darstellung verzichtet.

Fachlehrer darstellt. Dieser Blick nach außen ist grundsätzlich eine wichtige Facette externer Rückmeldungen, weil er von Lehrpersonen selbst schwer geleistet werden kann. Trotzdem kann argumentiert werden, dass ein objektiver Blick in die Unterrichtsgruppe in Kombination mit den subjektiven Einschätzungen der Lehrkraft zu weiteren Erkenntnissen beitragen würde: So könnte sich die Fachlehrerin/der Fachlehrer fragen, in welchen Inhalts- oder Handlungsbereichen ihre/seine Schüler/innen die höchsten Kompetenzen aufweisen. Der Vergleich zwischen den Werten der verschiedenen Kompetenzbereiche ist aber, wie oben erläutert, mit der derzeit eingesetzten Methodik schwierig. In der Rückmeldepraxis könnte auch die Darstellung der Ergebnisse auf einer 500er-Metrik eine Hürde für Lehrpersonen bedeuten. Während in Schularbeiten Ergebnisse mithilfe von Summenscores (bspw. 14 von 20 Punkten) berichtet werden, könnte die 500er-Metrik ungewohnt erscheinen und somit die Ableitung von Entwicklungspotenzial für den Unterricht erschweren.

Insgesamt scheint bei der derzeitigen Rückmeldeform der Mathematikkompetenzen auf Unterrichtsebene mit dem Ziel der Unterrichtsentwicklung Entwicklungspotenzial zu bestehen. Aus diesem Grund wollen wir im nächsten Abschnitt eine Darstellungsform diskutieren, die dieses Potenzial aufgreift. Die Vorschläge können eine Diskussionsgrundlage für die Weiterentwicklung von Rückmeldungen aus standardisierten Kompetenzmessungen mit dem Ziel der Unterrichtsentwicklung bilden. Für die vorgeschlagene weiterentwickelte Rückmeldeform muss das bestehende Testdesign nicht verändert werden.

### 10.3 Eine Option zur Weiterentwicklung der Rückmeldung von Mathematikkompetenzen auf Ebene der Unterrichtsruppen

Abbildung 6 zeigt eine weiterentwickelte Form der Rückmeldung von Mathematikkompetenzwerten auf Unterrichtsebene. Zum Zwecke der Unterrichtsentwicklung können Lehrpersonen Informationen über die Anzahl ihrer Schüler/innen entnehmen, die in einem Kompetenzknoten<sup>7</sup>, d.h. einer Kombination aus einem Inhalts- und einem Handlungsbereich, ein bestimmtes, zu definierendes Kompetenzniveau erreicht haben. Aus den die Bildungsstandards beschreibenden Can-Do-Statements<sup>8</sup> für Mathematik für das Ende der achten Schulstufe kann eine Beschreibung der Knotenpunkte entnommen werden: So ist beispielsweise der Kompetenzknoten „Zahlen und Maße/Darstellen und Modellbilden“ durch folgende Can-Do-Statements beschrieben: Die Schülerinnen und Schüler können (1) gegebene arithmetische Sach-

7 Die Kompetenzknoten sind in den Überprüfungen der Bildungsstandards mit durchschnittlich vier Items abgedeckt. Templin und Bradshaw (2013) diskutieren die Zuverlässigkeit von Individualklassifikationen im Hinblick auf die notwendige Anzahl von Items. Unabhängig davon zeigen Templin und Bradshaw (2013), dass die Individualklassifikationen aus kognitiv diagnostischen Modellen eine höhere Reliabilität als aus vergleichbaren mehrdimensionalen IRT-Modellen besitzen. Aggregiert man die individuellen Klassifikationen auf Gruppenebene, so ist dementsprechend auch auf der Gruppenebene von einer höheren Reliabilität im Vergleich zu vergleichbaren IRT-Modellen auszugehen.

8 <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20006166>



verhalte in eine (andere) mathematische Darstellung übertragen, wobei dafür das unmittelbare Einsetzen von Grundkenntnissen erforderlich ist, (2) gegebene arithmetische Sachverhalte in eine (andere) mathematische Darstellung übertragen, wobei dafür auch Verbindungen zu anderen mathematischen Inhalten (Begriffen, Sätzen, Darstellungen) oder Tätigkeiten hergestellt werden müssen, und (3) Aussagen über die Angemessenheit sowie über Stärken und Schwächen verschiedener mathematischer Darstellungen (Modelle) arithmetischer Sachverhalte machen und bewerten. Die Rückmeldung auf den Kompetenzknoten würde der Fachlehrerin/dem Fachlehrer nicht nur Informationen über Ergebnisse ihres/seines Unterrichts in den isolierten Inhaltsbereichen geben, sondern darüber hinaus Stärken und Schwächen in den inhaltlichen Kompetenzbereichen in Kombination mit den Handlungsbereichen aufzeigen. So könnte die Lehrkraft beispielsweise durch die Rückmeldung auf eine besondere Stärke ihrer Schüler/innen im Inhaltsbereich *Variable, funktionale Abhängigkeiten* in Verbindung mit dem Handlungsbereich *Interpretieren* hingewiesen werden und sie hätte die Möglichkeit, den eigenen Unterricht vor diesem Hintergrund zu reflektieren.

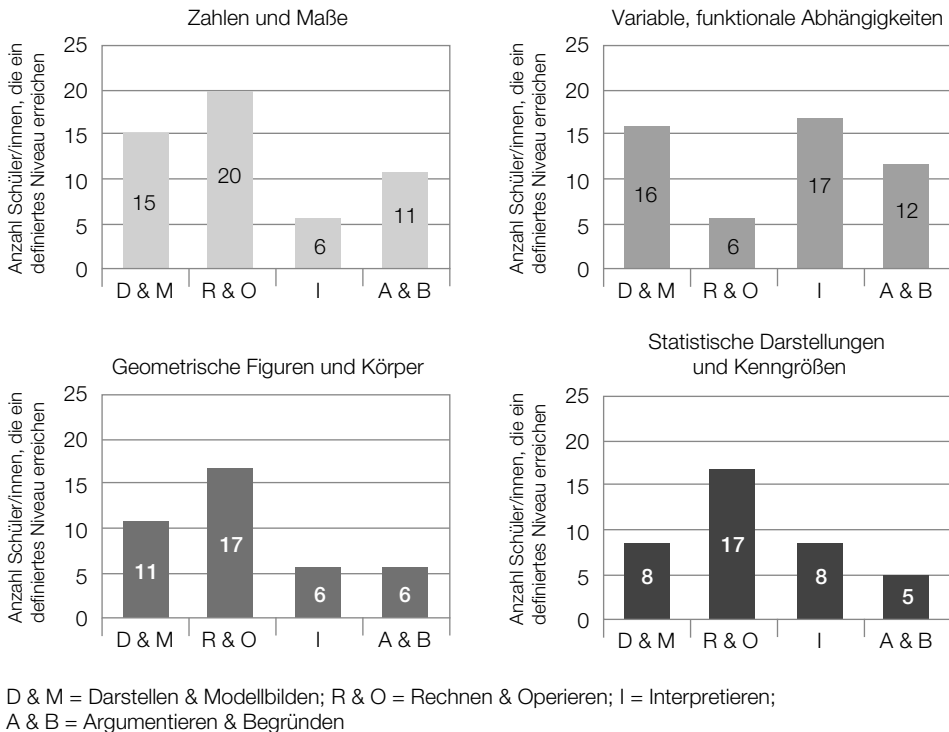


Abbildung 6: Weiterentwickelte Rückmeldung für die Ebene der Unterrichtsruppen auf Basis der Kombinationen zwischen Inhalts- und Handlungsbereichen.

Zur Erstellung dieser Rückmeldeform werden die codierten Testaufgaben aller Schüler/innen Österreichs mit einem kognitiv-diagnostischen Modell (CDM; als Überblick siehe DiBello, Roussos & Stout, 2007; Rupp, Leucht & Hartung, 2006) skaliert (George & Robitzsch, 2018). Dem Modell liegt eine Zuweisung (sog. Q-Matrix) der Testauf-

gaben zu genau einem der 16 Kompetenzknoten zugrunde, die bereits beim Erstellen der Aufgaben vorgenommen wird. Aus den Modellergebnissen werden für jede Schülerin bzw. jeden Schüler Individualklassifikationen ermittelt. Diese bestimmen, ob eine Schülerin/ein Schüler im jeweiligen Kompetenzknoten ein zu definierendes Kompetenzniveau erreicht hat oder nicht. Für die Unterrichtsruppenebene können die individuellen Klassifikationen aller Schüler/innen einer Unterrichtsruppe aggregiert werden. Durch diese Vorgehensweise liegen die Informationen über die Schüler/innen auf einer Skala vor, die hinsichtlich des absoluten Kriteriums „Beherrschung eines definierten Niveaus in einem Kompetenzknoten“ interpretiert werden kann. Daher erhalten Lehrkräfte Informationen, wie viele ihrer Schüler/innen bestimmte Kompetenzknoten auf einem bestimmten Niveau beherrschen. Diese weiterentwickelte Rückmeldemethode erlaubt auch einen Vergleich zwischen den Ausprägungen in den Kompetenzknoten (d.h. Aussagen wie „mehr Schüler/innen haben ein definiertes Niveau in *Interpretieren* im inhaltlichen Bereich *Variable, funktionale Abhängigkeiten* erreicht als im inhaltlichen Bereich *Statistische Darstellungen und Kenngrößen*“ sind möglich).

Das aktuelle Setting der Bildungsstandardüberprüfungen und deren Rückmeldungen zielt auf die Unterstützung von Qualitätsentwicklung auf der Ebene von Schulstandorten und Unterricht ab. Eine auf den Rückmeldungen basierende individuelle Förderung ist aufgrund der Überprüfung jeweils an den Schnittstellen des Schulsystems sowie der zeitlich um sechs bis acht Monate verschobenen Rückmeldung nicht möglich. Diese steht aber, wie oben betont, im aktuellen Setting nicht im Fokus. Die in Abbildung 4 präsentierte Individualrückmeldung dient daher ausschließlich der Information der Schüler/innen (und deren Eltern). Die jeweiligen Lehrer/innen erhalten, so wie es das derzeitige Gesetz vorzieht, nur anonymisierte Rückmeldungen (siehe Abbildung 3), die keinen direkten Schluss auf die individuellen Schüler/innen und deren Kompetenzen ermöglichen. In einem anderen Setting als dem aktuellen könnte aber ein verstärktes Interesse an Individualrückmeldungen bestehen.

Zu diesem Zweck könnte die in Abbildung 6 präsentierte weiterentwickelte Rückmeldung auch in nicht aggregierter Form als Individualklassifikationen angeboten werden (Abbildung 7). Hier kann die Lehrperson (oder der Schüler/die Schülerin und seine/ihre Eltern) direkt ablesen, in welchen der 16 Kompetenzknoten ein/e Schüler/in die angestrebten Kompetenzen besitzt und in welchen Förderbedarf besteht (vgl. Rupp, Templin & Henson, 2010, S. 65). Auch können Gruppen von Schülerinnen und Schülern mit gleichen Mustern von beherrschten Kompetenzen abgeleitet werden. Dies könnte als Grundlage für gemeinsame Fördermaßnahmen (z. B. auch für Begabtenförderung) eingesetzt werden.

	Zahlen und Maße					Variable, funktionale Abhängigkeit					Geometrische Figuren und Körper					Statistische Darstellungen und Kenngrößen				
	Darstellen & Modellbilden	Rechnen & Operieren	Interpretieren	Argumentieren & Begründen		Darstellen & Modellbilden	Rechnen & Operieren	Interpretieren	Argumentieren & Begründen		Darstellen & Modellbilden	Rechnen & Operieren	Interpretieren	Argumentieren & Begründen		Darstellen & Modellbilden	Rechnen & Operieren	Interpretieren	Argumentieren & Begründen	
Schüler 1	✓	✓	✗	✗		✓	✗	✓	✓		✓	✓	✓	✗		✓	✓	✓	✗	
Schüler 2	✓	✓	✗	✓		✓	✗	✗	✓		✓	✓	✗	✗		✓	✓	✓	✗	
	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	

Abbildung 7: Individuelle Profile von Schülerinnen bzw. Schülern über das Erreichen eines definierten Niveaus in den Kombinationen zwischen Inhalts- und Handlungsbereichen.

## 10.4 Diskussion

Im vorliegenden Kapitel wurden Entwicklungsdesiderata in der Rückmeldung von Mathematikkompetenzen auf Unterrichtsebene herausgestellt. Durch unterschiedliche Ziele der Rückmeldungen auf System- bzw. Unterrichtsebene unterscheiden sich auch die Anforderungen an die Rückmeldungen auf diesen beiden Ebenen. Auf Ebene der Unterrichtsebene steht die Unterstützung der Unterrichtsentwicklung im Mittelpunkt.

Im vorliegenden Kapitel wurde eine Option für eine weiterentwickelte Form der Rückmeldung auf Unterrichtsebene präsentiert (Abbildung 6), die das herausgestellte Entwicklungspotenzial aufgreift. Zum einen wurde die Detailtiefe der Rückmeldung erhöht und die Form an die Unterrichtspraxis in den Kompetenzknoten zwischen Inhalts- und Handlungsbereichen (statt in isolierten Kompetenzen) angepasst. Dies kann die Reflexion der Unterrichtspraxis durch die Lehrkraft unterstützen. Zum anderen erlaubt die weiterentwickelte Form einen direkten Vergleich zwischen den Kompetenzwerten der Schüler/innen in den 16 Kompetenzknoten. Durch die neue Form der Darstellungen werden anstelle von Mittelwerten in den Kompetenzbereichen auf der 500er-Metrik (gemeinsam mit einer inhaltlichen Einordnung durch den kriterialen Vergleich auf der Gesamtskala) nun absolute Anzahlen an Schülerinnen und Schülern (bzw. Prozentanteile) berichtet, die ein bestimmtes Kompetenzniveau je Kompetenzknoten erreichen.

Das vorliegende Kapitel liefert eine Grundlage für eine Diskussion möglicher Entwicklungsrichtungen von Rückmeldungen aus Kompetenzmessungen. Die aktuell bei Standardüberprüfungen eingesetzten Formen der Rückmeldung auf der einen Seite und die in diesem Kapitel aufgezeigten Möglichkeiten einer Weiterentwicklung auf der anderen Seite haben verschiedene Vor- und Nachteile. Diese sind mit Blick auf die Rahmenbedingungen, die Zielvorstellungen und die praktische Umsetzung zu diskutieren. Dabei scheinen zumindest folgende Aspekte von Relevanz:

- Der Bezug zwischen der jeweiligen Rückmeldung und dem Unterricht: Eine Rückmeldung nach Kompetenzknoten verkleinert die Unterschiede zwischen den Gliederungslogiken der Unterrichtspraxis und der Rückmeldung. In der weiterentwickelten Form der Rückmeldung treten die Kompetenzfacetten Inhalt und Handlung, so wie im Unterricht und in den Testitems, in Verknüpfung auf. In der aktuellen Form werden die Inhalts- und die Handlungsbereiche isoliert betrachtet. Mit der Weiterentwicklung könnte die gedankliche, reflexive Verbindung zwischen Rückmeldung und Unterricht, und somit die Ableitung von Maßnahmen aus der Rückmeldung, eventuell erleichtert werden.
- Der Detailgrad und die Skala der Rückmeldungen: Durch die getrennte Darstellung in den 16 Kompetenzknoten ist der inhaltliche Detaillierungsgrad der weiterentwickelten Rückmeldung größer als in der derzeit eingesetzten. Allerdings findet eine Einschränkung in Bezug auf den betrachteten Auflösungsgrad statt. Die Ergebnisse auf der 500er-Metrik decken ein breites Fähigkeitsspektrum ab und zeigen damit die Heterogenität der Kompetenzen innerhalb von Unterrichtsruppen und Schulen auf. Dagegen fokussiert die gezeigte Weiterentwicklungsoption auf ein (zu definierendes) Niveau, dessen Erreichen oder Nichterreichen berichtet wird. Auf der einen Seite ist in der weiterentwickelten Rückmeldung dadurch keine Information darüber zu entnehmen, wie weit Schüler/innen vom relevanten Schwellenwert entfernt sind. Auf der anderen Seite vereinfacht sich die Darstellung der Rückmeldung in der weiterentwickelten Form und die eingesetzten größeren Skalen können bei einem Fokus auf die Individualebene eine geringere statistische Unsicherheit der Kompetenzwerte aufweisen.
- Konzeption des Berichterstattungs- und Rückmeldesystems: In der aktuellen Form sind die Rückmeldungen an alle Ebenen in Bezug auf den Rückmeldezeitpunkt als auch in Bezug auf die Darstellungsformen und Bezugsnormen parallelisiert. Dies unterstützt ein durchgehendes gemeinsames Verständnis von Zielen, Überprüfung und Qualitätsentwicklung. Größere Anpassungen der Rückmeldeinformationen an die spezifischen Anforderungen einer Zielgruppe (etwa in diesem Fall der Fachlehrpersonen) sind vor diesem Hintergrund abzuwägen.

Als Ausblick dieser Diskussion sollten noch folgende Grundgedanken ergänzt werden: Neben den präsentierten Ausschnitten (Abbildung 2 bis 4) umfasst das aktuelle Rückmeldesystem als Interpretationsrahmen die kriteriale Sicht in Form der Kompetenzstufen (allerdings in Mathematik nur in Bezug auf die Gesamtskala) sowie eine soziale Bezugsnorm (als Vergleich zum Österreichwert bzw. zur eigenen Schule) und einen „fairen Vergleich“ (als Vergleich mit anderen Unterrichtsruppen mit ähnlichen Rahmenbedingungen). Überlegungen zur Weiterentwicklung der Unterrichtsruppenrückmeldung sollten diese Aspekte berücksichtigen. Daraus könnte gefolgert werden, dass (a) wesentliche Teile der aktuellen Rückmeldung weitergeführt und durch Informationen auf Basis der in diesem Kapitel vorgestellten Weiterentwicklung ergänzt werden oder (b) in einer weiteren Entwicklung dieser neu präsentierten Form der Unterrichtsruppenrückmeldungen Vergleichswerte zum österreichischen Mittelwert oder zu Kompetenzwerteprofilen von Unterrichtsruppen mit ähnlicher Kompo-

sition konzipiert werden. Weiters sind Studien notwendig, die untersuchen, ob Lehrkräfte diese weiterentwickelte Form der Rückmeldung als nützlich erachten und wie sie damit in der Praxis arbeiten. Interessant wäre in diesem Zusammenhang besonders, ob auf Basis der neuen Informationen eher als bisher konkrete Handlungen der Unterrichtsentwicklung abgeleitet werden.

## Literatur

- Blum, W. (2012). Die Bildungsstandards Mathematik. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret*. Berlin: Cornelsen.
- DiBello, L., Roussos, L. & Stout, W. (2007). Review of cognitively diagnostic assessment and a summary of psychometric models. In C. R. Rao & S. Sinharay (Hrsg.), *Handbook of Statistics, Volume 26, Psychometrics* (S. 979–1030). Amsterdam: Elsevier.
- George, A. C. & Robitzsch, A. (2018). Focusing on interactions between content and cognition: A new perspective on gender differences in mathematical sub-competencies. *Applied Measurement in Education*, 31, 79–97.
- Itzlinger-Bruneforth, U., Kuhn, J. T. & Kiefer, T. (2016). Testkonstruktion. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R* (S. 21–50). Wien: facultas.
- Luger-Bazinger, C., Freunberger, R. & Itzlinger-Bruneforth, U. (2016). Standard-Setting. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R* (S. 83–109). Wien: facultas.
- Pellegrino, J. W., Chudowsky, N. & Glaser, R. (Hrsg.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington: National Academy Press.
- Rupp, A. A., Leucht, M. & Hartung, R. (2006). Die Kompetenzbrille aufsetzen. Verfahren zur multiplen Klassifikation von Lernenden für Kompetenzdiagnostik in Unterricht und Testung. *Unterrichtswissenschaft*, 34, 195–219.
- Rupp, A. A., Templin, J. & Henson, R. A. (2010). *Diagnostic Measurement*. New York: The Guilford Press.
- Schreiner, C. & Breit, S. (Hrsg.). (2012). *Standardüberprüfung 2012. Mathematik, 8. Schulstufe. Bundesergebnisbericht*. Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens. Verfügbar unter [https://www.bifie.at/wp-content/uploads/2017/05/BiSt-UE\\_M8\\_2012\\_Bundesergebnisbericht.pdf](https://www.bifie.at/wp-content/uploads/2017/05/BiSt-UE_M8_2012_Bundesergebnisbericht.pdf)
- Schreiner, C. & Breit, S. (Hrsg.). (2014). *Standardüberprüfung 2013. Mathematik, 4. Schulstufe. Bundesergebnisbericht*. Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens. Verfügbar unter [https://www.bifie.at/wp-content/uploads/2017/05/BiSt-UE\\_M4\\_2013\\_Bundesergebnisbericht.pdf](https://www.bifie.at/wp-content/uploads/2017/05/BiSt-UE_M4_2013_Bundesergebnisbericht.pdf)
- Schreiner, C., Breit, S., Pointinger, M., Pacher, K., Neubacher, M. & Wiesner, C. (Hrsg.). (2018). *Standardüberprüfung 2017. Mathematik, 8. Schulstufe. Bundesergebnisbericht*. Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens. Verfügbar unter [https://www.bifie.at/wp-content/uploads/2019/02/BiSt-UE\\_M4\\_2018\\_Bundesergebnisbericht.pdf](https://www.bifie.at/wp-content/uploads/2019/02/BiSt-UE_M4_2018_Bundesergebnisbericht.pdf)
- Templin, J. & Bradshaw, L. (2013). Measuring the reliability of diagnostic classification model examinee estimates. *Journal of Classification*, 30, 251–275.
- Trendtel, M., Pham, G. & Yanagida, T. (2016). Skalierung und Linking. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R* (S. 185–224). Wien: facultas.

Wiesner, C., Schreiner, C., Breit, S., Kemethofer, D., George, A. C. & Angerer, S. (2016). Die Bedeutsamkeit der professionellen Reflexion und Rückmeldekultur für eine evidenzorientierte Schulentwicklung durch Bildungsstandardüberprüfungen. *Journal für Schulentwicklung*, 20, 18–26.